

WILLKIE FARR & GALLAGHER LLP

Benedict Y. Hur (SBN: 224018)
Simona Agnolucci (SBN: 246943)
Eduardo E. Santacana (SBN: 281668)
Lori C. Arakaki (SBN: 315119)
One Front Street, 34th Floor
San Francisco, CA 94111
Telephone: (415) 858-7400
Facsimile: (415) 858-7599
bhur@willkie.com
sagnolucci@willkie.com
esantacana@willkie.com
larakaki@willkie.com

Attorneys for
GOOGLE LLC

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO**

ANIBAL RODRIGUEZ, *et al.*, individually and on
behalf of all other similarly situated,

Plaintiffs,

vs.

GOOGLE LLC,

Defendant.

Case No. 3:20-CV-04688 RS

**DECLARATION OF STEVE GANEM IN
SUPPORT OF JOINT LETTER BRIEF
RE: GOOGLE PRESERVATION**

Judge: Hon. Alex G. Tse

47941117.1

1 I, Steve Ganem, declare as follows:

2 1. I am a Senior Product Manager at Google LLC (“Google”) with supervisory
3 authority concerning Google Analytics for Firebase (“GA4F”). Unless otherwise stated, the facts
4 I set forth in this declaration are based on my personal knowledge or knowledge I obtained
5 through my review of corporate records or other investigation. If called to testify as a witness, I
6 could and would testify competently to such facts under oath

7 2. I submit this declaration in support of Plaintiffs’ and Google’s Joint Letter Brief re:
8 Google Preservation.

9 3. Preserving the data that Plaintiffs are requesting—data from all users of all apps that
10 use GA4F—would be incredibly burdensome and costly. Doing so would require collecting from
11 the app measurement logs, which record 135 billion entries per day, on average.

12 4. For context, just 56 days’ worth of data totals over 27 petabytes of data per day (one
13 petabyte is ~one million gigabytes). It can take days or weeks of engineer and computer time for
14 even narrow queries to these logs to yield results.

15 5. Using public storage cost figures as an example, storing two years’ worth of such
16 data would cost \$3 million, and three years would cost \$6 million; preserving it indefinitely would
17 continue to increase cumulatively over time.

18 6. To illustrate the storage-related burden, a typical 40-server rack cabinet 7 feet tall
19 might hold 600 terabytes of data in it. To store fifty-six days’ worth of raw GA4F event-level
20 data, Google would need 23 full size server rack cabinets. To store one year’s worth of data,
21 Google would need 156 cabinets; for three years: 456 cabinets. That does not include the engineer
22 time required to query the data to produce a copy of it, which would also be extensive.

23 7. Google Analytics identifiers are, by design, pseudonymous to Google, in that they
24 are unique to the visitor's session on the particular customer's property, and are free of personally
25 identifiable information. The logs where data associated with these identifiers is stored similarly
26 cannot be used to identify a specific user, and Google does not log its own identifiers along
27 customers’ Google Analytics identifiers. This means that the data in the logs is edited by Google
28

47941117.1

1 at the time that Google receives it, before it is saved to non-ephemeral memory, to ensure that a
2 particular Google user cannot be identified from the entries in the This log editing includes
3 sophisticated encryption techniques in concert with redaction, making it virtually impossible
4 for anyone at Google to re-join entries in these pseudonymous logs with a user's identity. This
5 anonymization process is a core part of how Google honors user transparency and control. To
6 illustrate, I received from counsel a list of AdIDs provided by Plaintiffs' counsel taken from the
7 devices owned by certain Plaintiffs. These AdIDs identify a particular device. Upon investigating
8 how to identify which entries in the GA for Firebase logs belong to those AdIDs, my team and I
9 determined that it would take days or weeks of effort by several engineers to identify and produce
10 GA for Firebase data associated with AdIDs. Further, if Plaintiffs had not provided those AdIDs,
11 Google would not have a way of knowing them because those IDs are not stored alongside users'
12 identities. The difficulty Google would have in matching the AdIDs to entries in the logs is by
13 design; it ensures that Google's prohibition on the "joining" of pseudonymous entries with a
14 user's identity without user consent is effective.

15 I declare under the penalty of perjury that the foregoing is true and correct.

16
17 Executed on November 10, 2021.

18
19 
20 STEVE GANEM
21
22
23
24
25
26
27
28